Volume 202, number 1 FEBS 3717 June 1986

Measles virus hemagglutinin

Removal of the initiator methionine in the mature protein, and evidence for further processing to produce a 'ragged' end

Tamas M. Varsanyi, Gunilla Lundquist, Erling Norrby and Hans Jörnvall

Departments of Virology and Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden

Received 15 April 1986

Measles virus hemagglutinin has been isolated by immunoadsorption. The total composition of the protein and its N-terminal amino acid sequence give data matching the structure indirectly deduced from cDNA. However, direct analysis of the hemagglutinin also shows that the mature protein is proteolytically processed and has a partly heterogeneous N-terminus. The initiator Met is removed, and non-stoichiometrically also the second residue.

Protein processing *Measles virus* *Hemagglutinin* *Amino acid sequence* *Sequence homology*

1. INTRODUCTION

The envelope of measles virus consists of a lipid bilayer of cellular origin into which the viral glycoproteins, the hemagglutinin and the fusion protein, are inserted as integral membrane proteins. The measles virus hemagglutinin protein chain has an estimated M_r of 79000–80000 [1–3], and is responsible for virus adsorption to cells by interaction with a cell receptor [4]. The hemagglutinin molecule is a disulfide-bonded structure, probably a dimer of a single type of polypeptide chain [5,6], forming a spike with truncated conical shape [7] on the surface of virions and infected cells.

Most viral membrane-associated glycoproteins analyzed have membrane-inserted segments in the C-terminal regions, and signal peptides as well as initiator methionine residues that are cleaved off in post-translational events [8–10]. Unlike this common orientation of membrane glycoproteins, influenza A virus neuraminidase has been suggested, from analysis of a corresponding cDNA, to be attached to the viral membrane by its N-terminal region [11], for which a non-processed N-terminal end, absence of a signal peptide segment, and presence of the initiator methionine have been shown by direct protein analysis [12]. A similar structural organization was concluded for the influenza B neuraminidase [13]. The N-terminus of hemagglutinin-neuraminidase protein of Newcastle disease virus appears to be blocked, and is assumed to be inserted into the membrane [10]. cDNA-deduced amino acid sequences of the hemagglutinin-neuraminidase proteins of Sendai virus [14] and Simian virus 5 [15] reveal one major hydrophobic region near the N-terminus assumed to represent the trans-membraneous part. The similarly deduced 617-residue amino acid sequence of the measles virus hemagglutinin also has a major hydrophobic region at positions 35–58, in the N-terminal segment of the molecule [16]. However, the measles virus hemagglutinin protein has not been directly analyzed and the aim of the present study was to demonstrate whether the measles virus hemagglutinin has an unprocessed N-terminal end or if it undergoes post-translational cleavage.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

2. MATERIALS AND METHODS

2.1. *Isolation of the measles virus hemagglutinin glycoprotein*

The protein was isolated by affinity chromatography as a major step [17]. Confluent HeLa cells were infected with plaque purified LEC (KI) strain of measles virus. At full infection, cells were pelleted and lysed by addition of 0.01 M Tris-HCl (pH 7.8), 0.15 M NaCl, 1% aprotinin, 2 mM phenylmethylsulfonyl fluoride, 5 mM EDTA and 3% Triton X-100. From the clarified solution, the hemagglutinin was extracted by anti-hemagglutinin monoclonal antibodies immobilized on protein A-Sepharose. The immunocomplex was excessively washed with lysis buffer and the hemagglutinin was separated from immunoglobulins by preparative polyacrylamide gel electrophoresis in the presence of mercaptoethanol. Polypeptide bands, identified as the 79 kDa hemagglutinin monomer, were excised from the gels, and the protein extracted in 1% SDS by rocking at 37°C for 72 h. The hemagglutinin was then precipitated with trichloroacetic acid, washed with acetone and subjected to further analysis.

2.2. *Structural analysis*

Amino acid compositions were determined by acid hydrolysis and subsequent amino acid analysis on a Beckman 121M instrument. Gas-phase sequencer degradations were carried out in an Applied Biosystems 790A instrument. Phenylthiohydantoin derivatives were identified by high-performance liquid chromatography (Hewlett Packard 1090B).

3. RESULTS

3.1. *Total composition*

The total composition of the hemagglutinin as determined by acid hydrolysis is shown in table 1. The composition of the preparation analyzed is in agreement with that indirectly determined from analysis of the corresponding cDNA, as also shown in table 1. Consequently, correct assignments in the indirectly derived structure are supported.

Table 1

Amino acid composition of measles virus hemagglutinin protein

Residue	Determined (mol%)	Predicted [16] (mol%)
Asx	10.5	10.8
Thr	5.9	6.1
Ser	8.3	8.2
Glx	9.1	8.2
Pro	7.6	5.7
Gly	8.5	6.7
Ala	5.0	5.0
Val	7.3	8.1
Met	2.1	2.7
Ile	4.7	6.1
Leu	10.5	11.4
Tyr	4.3	4.4
Phe	3.6	3.7
Lys	5.6	5.1
His	2.1	2.5
Arg	5.0	5.2

Cys and Trp presently not analyzed and not included in the values above

3.2. *N-terminal sequence analysis*

Results of gas-phase sequencer analysis of the protein are given in table 2. One major sequence was followed for 15 cycles. It corresponds to residue 2–16 of the cDNA-derived indirectly deduced amino acid sequence for the hemagglutinin [16], establishing that the major form of the mature protein has lost the initiator methionine.

In addition, a second sequence was clearly visible, yielding the same structure, but one cycle earlier in each case (table 2). It corresponds to residues 3–17 in the cDNA-deduced amino acid sequence [16], showing that N-terminal proteolytic processing has progressed one step further in a substantial part of the mature polypeptide molecules. The amounts of the 'longer' and 'shorter' polypeptide chains were roughly 75 and 25%, respectively (table 2), yielding a 'ragged' N-terminal end. Together, the recovery of these two structures account for about half the molar yield in relation to the amount of protein initially applied to the glass fiber in the gas-phase sequencer. Considering that yields of initial coupling frequently are only somewhat over 50%, the present

Table 2
Results of amino acid sequence analysis

Cycle	Major yield (pmol)	Minor yield (% of total)
1	Ser 200	Pro 24
2	Pro 302	Gln 19
3	Gln 324	Arg 26
4	Arg 64	Asp 25
5	Asp 162	Arg 24
6	Arg 82	Ile 22
7	Ile 374	Asn 24
8	Asn 155	Ala 17
9	Ala 253	Phe 20
10	Phe 384	Tyr 25
11	Tyr 259	Lys 25
12	Lys 222	Asp 30
13	Asp 118	Asn 23
14	Asn 111	Pro 22
15	Pro 198	His 25
Average yield of total	77%	23%

As shown, two residues were detected in each cycle but could be distinguished because of differences in yield, the major peptide representing about 80% of the total recovery, and the minor about 20%. Results were obtained by degradation of about 1 nmol hemagglutinin protein and phenylthiohydantoin identification by reverse-phase high-performance liquid chromatography

recoveries show that the major structure determined is also the major polypeptide chain present in the hemagglutinin preparation.

Further structures in minor yield were also visible especially in the initial cycles of the degradations, and are compatible with a low amount of additional polypeptide chains. However, they could not be identified as further processing variants of the hemagglutinin molecule, and might therefore represent contaminants or completely different cleavage products. Degradations of one further preparation also showed the presence of additional contaminants, while still another preparation of the hemagglutinin fully confirmed the present results, establishing the reproducibility.

4. DISCUSSION

4.1. *Nature of the measles virus hemagglutinin polypeptide*

The structures of the hemagglutinin polypeptide from measles virus and hemagglutinin-neuraminidase proteins from Simian virus 5 and Sendai virus have been predicted from analyses of corresponding cDNAs [14–16]. However, the measles virus protein itself has never been directly analyzed previously. The total composition (table 1) and the N-terminal amino acid sequence (table 2) now obtained therefore form a direct support for the correctness of the hemagglutinin structure indirectly deduced. Furthermore, the protein data establish that the mature hemagglutinin is N-terminally processed by removal of the initiator methionine (formally shifting the numbering system of the hemagglutinin amino acid sequence

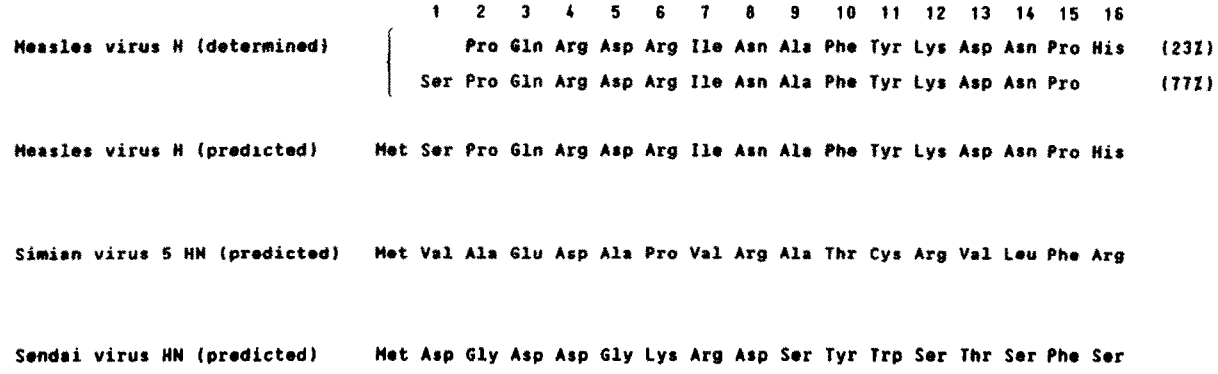

Fig.1. Comparison of the N-terminal structures of three paramyxovirus hemagglutinin (H) and hemagglutinin-neuraminidase (HN) proteins. The directly determined data for the measles virus hemagglutinin are from table 2, the predicted data from the report of the corresponding cloned cDNA [16]. Similarly, the predicted data for Simian virus 5 are from [15] and for Sendai virus hemagglutinin-neuraminidase from [14]. The numbering system is given for the major form of the mature measles virus protein now analyzed.

[16] by one residue). They also reveal that the polypeptide is sensitive to further processing and that a minor part of the mature protein consists of polypeptides starting at position 2 (corresponding to position 3 of the cDNA-deduced amino acid sequence [16]), giving a heterogeneous N-terminal starting position in the mature preparation. Consequently, measles virus hemagglutinin belongs to the proteins with a ragged N-terminal end. Similar heterogeneity in N-terminal starting position is also observed in mitochondrial proteins [18], where the ragged N-terminal ends probably reflect native processing differences. Ragged N-terminal ends are also common in several enzymes and hormones, where the origin of the heterogeneity is still not fully established (e.g. [19–21]).

The N-terminal structure is of interest in relation to acetylation of polypeptide chains. Thus, structural polypeptides are frequently N-terminally acetylated [22], the acetylation being especially common with N-terminal serine, which is the residue now detected after the initial removal of the initiator methionine (table 2), as well as with methionine itself [22]. Consequently, N-terminal acetylation with resulting blocked polypeptide chains could have been expected from the structure of the cDNA-deduced amino acid sequence [16], preventing direct N-terminal analysis [23], as in many other indirectly deduced structures. Although a small proportion of *N*-acetyl-Met blocked or *N*-acetyl-Ser blocked polypeptide chains in the hemagglutinin preparation cannot be excluded, the present yields upon direct sequence analysis (table 2) show that the polypeptide chains with free N-terminal serine are the major constituents (table 2). In fact, empirical rules for the methionine removal of nascent proteins in general [24] also predict removal of the initiator methionine when the subsequent residue is serine.

Finally, in relation to N-terminal segments of the hemagglutinin and the fusion proteins, it is of interest to note a difference in pattern for the variability of special regions in the two paramyxovirus surface proteins. Thus, comparison of the F_1 subunit of the fusion protein of the measles virus with other paramyxovirus fusion proteins shows extreme conservation in the N-terminal end [17], while the membrane-inserted N-terminal region of the measles virus hemagglutinin now analyzed exhibits extensive variability in comparison to corresponding regions of other paramyxoviruses (fig.1). In biological terms, these findings indicate that for the membrane fusion process a strictly conserved region is a prerequisite, while the events in which the hemagglutinin on the cytoplasmic side of the membrane participates (i.e. presumably in matrix-nucleocapsid protein recognition in the budding process [8]) apparently permit large variations.

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Medical Research Council (project 16X-00116 and 03X-3532) and the Swedish Cancer Society (project 1806).

REFERENCES

[1] Graves, M.C., Silver, S.M. and Choppin, P.W. (1978) Virology 88, 254–263.
[2] Hardwick, J.M. and Bussell, R.H. (1978) J. Virol. 25, 687–692.
[3] Tyrrell, D.L.J. and Norrby, E. (1978) J. Gen. Virol. 39, 219–229.
[4] Fenger, T. and Howe, C. (1979) Proc. Soc. Exp. Biol. Med. 162, 299–303.
[5] Casali, P., Sissons, J.G.P., Fujinami, R.S. and Oldstone, M.B.A. (1981) J. Gen. Virol. 54, 161–171.
[6] Graves, M.C. (1981) J. Virol. 38, 224–230.
[7] Varsanyi, T.M., Utter, G. and Norrby, E. (1984) J. Gen. Virol. 65, 355–366.
[8] Wiley, D.C. (1985) in: Virology (Fields, B.N. et al. eds) pp.45–67, Raven, New York.
[9] Wiley, D.C. (1983) Nature 303, 19–20.
[10] Schuy, W., Garten, W., Linder, D. and Klenk, H.-D. (1984) Virus Res. 1, 415–426.
[11] Fields, J., Winter, G. and Brownlee, G.G. (1981) Nature 290, 213–217.
[12] Blok, J., Air, G.M., Laver, W.G., Ward, C.W., Lilley, G.G., Woods, E.F., Roxburgh, C.M. and Inglis, A.S. (1982) Virology 119, 109–121.
[13] Shaw, M.W., Lamb, R.A., Erickson, B.W., Briedis, D.J. and Choppin, P.W. (1982) Proc. Natl. Acad. Sci. USA 79, 6817–6821.
[14] Blumberg, B., Giorgi, C., Roux, L., Raju, R., Dowling, P., Chollet, A. and Kolakofsky, D. (1985) Cell 41, 269–278.
[15] Hiebert, S.W., Paterson, R.G. and Lamb, R.A. (1985) J. Virol. 54, 1–6.

[16] Alkhatib, G. and Briedis, D.J. (1986) Virology, 150, 479–490.
[17] Varsanyi, T.M., Jörnvall, H. and Norrby, E. (1985) Virology 147, 110–117.
[18] Hempel, J., Kaiser, R. and Jörnvall, H. (1985) Eur. J. Biochem. 153, 13–28.
[19] Jörnvall, H., Carlquist, M., Kwauk, S., Otte, S.C., McIntosh, C.H.S., Brown, J.C. and Mutt, V. (1981) FEBS Lett. 123, 205–210.
[20] Sara, V.R., Carlsson-Skwirut, C., Andersson, C., Hall, E., Sjögren, B., Holmgren, A. and Jörnvall, H. (1986) Proc. Natl. Acad. Sci. USA, in press.
[21] Jörnvall, H., Pohl, G., Bergsdorf, N. and Wallén, P. (1983) FEBS Lett. 156, 47–50.
[22] Persson, B., Flinta, C., Von Heijne, G. and Jörnvall, H. (1985) Eur. J. Biochem. 152, 523–527.
[23] Ålin, P., Mannervik, B. and Jörnvall, H. (1985) FEBS Lett. 182, 319–322.
[24] Flinta, C., Persson, B., Jörnvall, H. and Von Heijne, G. (1986) Eur. J. Biochem. 154, 193–196.